Docket No.: 428.1056

UNITED STATES PATENT & TRADEMARK OFFICE

Examiner: Unknown	Art Unit: Unknown
Applicant:	KIM, Young Tae, et al.
Serial No.:	10/551,508
Filed	September 29, 2005
For:	**GENE INVOLVED IN THE BIOSYNTHESIS OF CAROTENOID AND MARINE MICROORGANISM, PARACOCCUS HAEUNDAENSIS, PRODUCING THE CAROTENOID**

RESPONSE TO NOTICE TO FILE MISSING PARTS

Commissioner for Patents   June 5, 2006
P.O. Box 1450
Alexandria, VA 22313-1450

Sir:

In response to the Notice to File Missing Parts ("Notice") dated May 10, 2006, enclosed herewith is an executed Declaration. As indicated in the Notice, the previous payment will be applied to the requisite surcharge of $65.00 (at Small Entity).

The Notice indicated that the originally submitted Computer Readable Form copy of the Sequence Listing does not comply with the requirements of 37 CFR 1.822 and/or 1.823, because the original copy contains informalities, for example, misaligned amino acid numbering.

A substitute Computer Readable copy of the Sequence Listing is enclosed herewith, together with the paper copy of the Sequence Listing. The misaligned amino acid numbering has been corrected. The substitute Sequence Listing also provides the application serial number and the filing date. No other changes have been made to the Sequence Listing, and the required Declaration to that effect is enclosed.
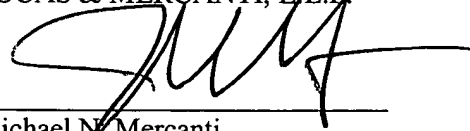
Pursuant to 37 C.F.R. §1.136(a)(3), please treat this and any concurrent or future reply in this application that requires a petition for an extension of time for its timely submission as incorporating a petition for extension of time for the appropriate length of time. The fee associated therewith is to be

1

charged to Deposit Account No. 02-2275.

Any deficiency or overpayment should be charged or credited to Deposit Account No. 02-2275. A duplicate copy of this sheet is enclosed.

Respectfully submitted,

LUCAS & MERCANTI, L.L.P.

Michael N. Mercanti
Reg. No. 33,966

LUCAS & MERCANTI, LLP
475 Park Avenue South
New York, New York 10016
Phone: 212-661-8000
Fax: 212-661-8002

UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

U.S. APPLICATION NUMBER NO.	FIRST NAMED APPLICANT	ATTY. DOCKET NO.
10/551,508	Young Tae KIM	428.1056

RECEIVED
WITH
THANKS

MAY 1 5 2006

Michael N Mercanti
475 Park Avenue South
New York, NY 10016

LUCAS & MERCANTI LLP

INTERNATIONAL APPLICATION NO.
PCT/KR04/00752

I.A. FILING DATE	PRIORITY DATE
03/31/2004	03/31/2003

CONFIRMATION NO. 8016
371 FORMALITIES LETTER

OC000000018713342

Date Mailed: 05/10/2006

NOTIFICATION OF MISSING REQUIREMENTS UNDER 35 U.S.C. 371 IN THE UNITED STATES DESIGNATED/ELECTED OFFICE (DO/EO/US)

The following items have been submitted by the applicant or the IB to the United States Patent and Trademark Office as a Designated / Elected Office (37 CFR 1.495).

- Indication of Small Entity Status
- Copy of the International Application filed on 09/29/2005
- Copy of the International Search Report filed on 09/29/2005
- Preliminary Amendments filed on 09/29/2005
- Biochemical Sequence Diskette filed on 09/29/2005
- Oath or Declaration filed on 09/29/2005
- Biochemical Sequence Listing filed on 09/29/2005
- Request for Immediate Examination filed on 09/29/2005
- U.S. Basic National Fees filed on 09/29/2005

The applicant needs to satisfy supplemental fees problems indicated below.

The following items **MUST** be furnished within the period set forth below in order to complete the requirements for acceptance under 35 U.S.C. 371:

- Oath or declaration of the inventors, in compliance with 37 CFR 1.497(a) and (b), identifying the application by the International application number and international filing date. The current oath or declaration does not comply with 37 CFR 1.497(a) and (b) in that it:
  - is not executed in accordance with either 37 CFR 1.66 or 37 CFR 1.68.

- To avoid abandonment, a surcharge (for late submission of filing fee, search fee, examination fee or oath or declaration) as set forth in 37 CFR 1.492(h) of $65 for a small entity in compliance with 37 CFR 1.27, must be submitted with the missing items identified in this letter.

SUMMARY OF FEES DUE:

Total additional fees required for this application is $-115 for a Small Entity:

- $65 Surcharge.

    - A copy of the "Sequence Listing" in computer readable form has been submitted. However, the content of the computer readable form does not comply with the requirements of 37 CFR 1.822 and/or 1.823, as indicated on the attached copy of the marked -up "Raw Sequence Listing." Applicant must provide a substitute computer readable form (CRF) copy of the "Sequence Listing" and a statement that the content of the sequence listing information recorded in computer readable form is identical to the written (on paper or compact disc) sequence listing and, where applicable, includes no new matter, as required by 37 CFR 1.821(e), 1.821(f), 1.821(g), 1.825(b), or 1.825(d).

        (A previous payment of $180 will be applied to the additional fees indicated above.)

Applicant is cautioned that correction of the above items may cause the specification and drawings page count to exceed 100 pages. If the specification and drawings exceed 100 pages, applicant will need to submit the required application size fee.

For questions regarding compliance to 37 CFR 1.821-1.825 requirements, please contact:

- **For Rules Interpretation, call (571) 272-0951**
- **For Patentin Software Program Help, call Patent EBC at 1-866-217-9197 or directly at 703-305-3028 / 703-308-6845 between the hours of 6 a.m. and 12 midnight, Monday through Friday, EST.**
- **Send e-mail correspondence for Patentin Software Program Help @ ebc@uspto.gov**

ALL OF THE ITEMS SET FORTH ABOVE MUST BE SUBMITTED WITHIN TWO (2) MONTHS FROM THE DATE OF THIS NOTICE OR BY 32 MONTHS FROM THE PRIORITY DATE FOR THE APPLICATION, WHICHEVER IS LATER. FAILURE TO PROPERLY RESPOND WILL RESULT IN ABANDONMENT.

The time period set above may be extended by filing a petition and fee for extension of time under the provisions of 37 CFR 1.136(a).

Applicant is reminded that any communications to the United States Patent and Trademark Office must be mailed to the address given in the heading and include the U.S. application no. shown above (37 CFR 1.5)

*A copy of this notice **MUST** be returned with the response.*

FRANCINE YOUNG

Telephone: (703) 308-9140 EXT 215

PART 1 - ATTORNEY/APPLICANT COPY

U.S. APPLICATION NUMBER NO.	INTERNATIONAL APPLICATION NO.	ATTY. DOCKET NO.
10/551,508	PCT/KR04/00752	428.1056

FORM PCT/DO/EO/905 (371 Formalities Notice)

RECEIVED
WITH
THANKS

RAW SEQUENCE LISTING
ERROR REPORT

MAY 15 2006

LUCAS & MERCANTI LLP

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) detected errors when processing the following computer readable form:

Application Serial Number: _10/551, 508_

Source: _PCT_

Date Processed by STIC: _10/14/2005_

THE ATTACHED PRINTOUT EXPLAINS DETECTED ERRORS.
PLEASE FORWARD THIS INFORMATION TO THE APPLICANT BY EITHER:
1) INCLUDING A COPY OF THIS PRINTOUT IN YOUR NEXT COMMUNICATION TO THE APPLICANT, WITH A NOTICE TO COMPLY or,
2) TELEPHONING APPLICANT AND FAXING A COPY OF THIS PRINTOUT, WITH A NOTICE TO COMPLY
FOR CRF SUBMISSION AND PATENTIN SOFTWARE QUESTIONS, PLEASE CONTACT MARK SPENCER, TELEPHONE: 571-272-2510; FAX: 571-273-0221

TO REDUCE ERRORED SEQUENCE LISTINGS, PLEASE USE THE **CHECKER VERSION 4.2.2 PROGRAM**, ACCESSIBLE THROUGH THE U.S. PATENT AND TRADEMARK OFFICE WEBSITE. SEE BELOW FOR ADDRESS:
       http://www.uspto.gov/web/offices/pac/checker/chkrnote.htm

Applicants submitting genetic sequence information electronically on diskette or CD-Rom should be aware that there is a possibility that the disk/CD-Rom may have been affected by treatment given to all incoming mail.
Please consider using alternate methods of submission for the disk/CD-Rom or replacement disk/CD-Rom.
Any reply including a sequence listing in electronic form should NOT be sent to the 20231 zip code address for the United States Patent and Trademark Office, and instead should be sent via the following to the indicated addresses:

1. EFS-Bio (**<http://www.uspto.gov/ebc/efs/downloads/documents.htm>** , **EFS Submission User Manual** - ePAVE)
2. U.S. Postal Service: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450
3. **Hand Carry, Federal Express, United Parcel Service, or other delivery service (EFFECTIVE 01/14/05):**
   U.S. Patent and Trademark Office, Mail Stop Sequence, Customer Window, Randolph Building, 401 Dulany Street, Alexandria, VA 22314

Revised 01/24/05

Raw Sequence Listing Error Summary

ERROR DETECTED	SUGGESTED CORRECTION	SERIAL NUMBER:

ATTN: NEW RULES CASES PLEASE DISREGARD ENGLISH ALPHA HEADERS, WHICH WERE INSERTED BY PTO SOFTWARE

1 ____ Wrapped Nucleics / Wrapped Aminos — The number/text at the end of each line "wrapped" down to the next line. This may occur if your file was retrieved in a word processor after creating it. Please adjust your right margin to :3; this will prevent "wrapping."

2 ____ Invalid Line Length — The rules require that a line not exceed 72 characters in length. This includes white spaces.

3 ____ Misaligned Amino Numbering — The numbering under each 5th amino acid is misaligned. Do not use tab codes between numbers; use space characters, instead.

4 ____ Non-ASCII — The submitted file was not saved in ASCII(DOS) text, as required by the Sequence Rules. Please ensure your subsequent submission is saved in ASCII text.

5 ____ Variable Length — Sequence(s) ____ contain n's or Xaa's representing more than one residue. Per Sequence Rules, each n or Xaa can only represent a single residue. Please present the maximum number of each residue having variable length and indicate in the <220>-<223> section that some may be missing.

6 ____ PatentIn 2.0 "bug" — A "bug" in PatentIn version 2.0 has caused the <220>-<223> section to be missing from amino acid sequences(s) _____. Normally, PatentIn would automatically generate this section from the previously coded nucleic acid sequence. Please manually copy the relevant <220>-<223> section to the subsequent amino acid sequence. **This applies to the mandatory <220>-<223> sections for Artificial or Unknown sequences.**

7 ____ Skipped Sequences (OLD RULES) — Sequence(s) ____ missing. If intentional, please insert the following lines for each skipped sequence:
(2) INFORMATION FOR SEQ ID NO:X: (insert SEQ ID NO where "X" is shown)
(i)   SEQUENCE CHARACTERISTICS: (Do not insert any subheadings under this heading)
(xi) SEQUENCE DESCRIPTION:SEQ ID NO:X: (insert SEQ ID NO where "X" is shown)
This sequence is intentionally skipped

Please also adjust the "(ii) NUMBER OF SEQUENCES:" response to include the skipped sequences.

8 ____ Skipped Sequences (NEW RULES) — Sequence(s) _____ missing. If intentional, please insert the following lines for each skipped sequence.
<210> sequence id number
<400> sequence id number
000

9 ____ Use of n's or Xaa's (NEW RULES) — Use of n's and/or Xaa's have been detected in the Sequence Listing.
Per 1.823 of Sequence Rules, use of <220>-<223> is MANDATORY if n's or Xaa's are present.
In <220> to <223> section, please explain location of n or Xaa, and which residue n or Xaa represents.

10 ____ Invalid <213> Response — Per 1.823 of Sequence Rules, the only valid <213> responses are: Unknown, Artificial Sequence, or scientific name (Genus/species). <220>-<223> section is required when <213> response is Unknown or is Artificial Sequence

11 ____ Use of <220> — Sequence(s) _____ missing the <220> "Feature" and associated numeric identifiers and responses. Use of <220> to <223> is MANDATORY if <213> "Organism" response is "Artificial Sequence" or "Unknown." Please explain source of genetic material in <220> to <223> section.
(See "Federal Register," 06/01/1998, Vol. 63, No. 104, pp. 29631-32)   (Sec. 1.823 of Sequence Rules)

12 ____ PatentIn 2.0 "bug" — Please do not use "Copy to Disk" function of PatentIn version 2.0. This causes a corrupted file, resulting in missing mandatory numeric identifiers and responses (as indicated on raw sequence listing). Instead, please use "File Manager" or any other manual means to copy file to floppy disk.

13 ____ Misuse of n/Xaa — "n" can only represent a single nucleotide; "Xaa" can only represent a single amino acid

AMC – Biotechnology Systems Branch – 09/09/2003

PCT

RAW SEQUENCE LISTING DATE: 10/14/2005
PATENT APPLICATION: US/10/551,508 TIME: 11:04:23

Input Set : N:\PANTU\10551508.raw.txt
Output Set: N:\CRF4\10142005\J551508.raw

```
    3 <110> APPLICANT: KIM, Young Tae
    4        LEE, Jae Hyung
    7 <120> TITLE OF INVENTION: Gene involved in the biosynthesis of carotenoid and marine
    8        microorganism, paracoccus haeundaesis, producing the
    9        carotenoid
   11 <130> FILE REFERENCE: 428.1056
C--> 13 <140> CURRENT APPLICATION NUMBER: US/10/551,508
C--> 13 <141> CURRENT FILING DATE: 2005-09-29
   13 <150> PRIOR APPLICATION NUMBER: PCT/KR2004/000752
   14 <151> PRIOR FILING DATE: 2003-03-31
   16 <150> PRIOR APPLICATION NUMBER: KR2003-20222
   17 <151> PRIOR FILING DATE: 2003-03-31
   19 <150> PRIOR APPLICATION NUMBER: KR2003-20023
   20 <151> PRIOR FILING DATE: 2003-03-31
   22 <160> NUMBER OF SEQ ID NOS: 18
   24 <170> SOFTWARE: KopatentIn 1.71
```

ERRORED SEQUENCES

```
   52 <210> SEQ ID NO: 3
   53 <211> LENGTH: 1454
   54 <212> TYPE: DNA
   55 <213> ORGANISM: Paracoccus haeundaesis
   57 <400> SEQUENCE: 3
   58 caacttgaga gtttgatcct ggctcagaac gaacgctggc ggcaggctta acacatgcaa     60
   60 gtcgagcgag accttcgggt ctagcggcgg acgggtgagt aacgcgtggg aacgtgccct    120
   62 tctctacgga atagccccgg gaaactggga gtaataccgt atacgccctt tggggaaag    180
   64 atttatcgga gaaggatcgg cccgcgttgg attaggtagt tggtggggta atggcccacc    240
   66 aagccgacga tccatagctg gtttgagagg atgatcagcc acactgggac tgagacacgg    300
   68 cccagactcc tacgggaggc agcagtgggg aatcttagac aatggggca accctgatct    360
   70 agccatgccg cgtgagtgat gaaggcctta gggttgtaaa gctctttcag ctgggaagat    420
   72 aatgacggta ccagcagaag aagccccggc taactccgtg ccagcagccg cggtaatacg    480
   74 gagggggcta gcgttgttcg gaattactgg gcgtaaagcg cacgtgggcg gactggaaag    540
   76 tcagaggtga atcccaggg ctcaaccttg gaactgcctt tgaaactatc agtctggagt    600
   78 tcgagagagg tgagtggaat tccgagtgta gaggtgaaat tcgtagatat tcggaggaac    660
   80 accagtggcg aaggcggctc actggctcga tactgacgct gaggtgcgaa agcgtgggga    720
   82 gcaaacagga ttagatacccc tggtagtcca cgccgtaaac gatgaatgcc agacgtcggc    780
   84 aagcatgctt gtcggtgtca cacctaacgg attaagcatt ccgcctgggg agtacggtcg    840
   86 caagattaaa actcaaagga attgacgggg gcccgcacaa gcggtggagc atgtggttta    900
   88 attcgaagca acgcgcagaa ccttaccaac ccttgacatg gcaggaccgc tggagagatt    960
   90 cagctttctc gtaagagacc tgcacacagg tgctgcatgg ctgtcgtcag ctcgtgtcgt   1020
   92 gagatgttcg gttaagtccg gcaacgagcg caacccacgt ccctagttgc cagcattcag   1080
```

RAW SEQUENCE LISTING   DATE: 10/14/2005
PATENT APPLICATION: US/10/551,508   TIME: 11:04:23

Input Set : N:\FANTU\10551508.raw.txt
Output Set: N:\CRF4\10142005\J551508.raw

```
 94 ttgggcactc tatggaaact gccgatgata agtcggagga aggtgtggat gacgtcaagt  1140
 96 tctcatggcc cttacgggtt gggctacaca cgtgctacaa tggtggtgac agtgggttaa  1200
 98 tccccaaaag ccatctcagt tcggattgtc ctctgcaact cgagggcatg aagttggaat  1260
100 cgctagtaat cgcggaacag catgccgcgg tgaatacgtt cccgggcctt gtacacaccg  1320
102 cccgtcacac catgggagtt ggttctaccc gacgacgctg cgctaacctt cgggggcag   1380
104 gcggccacgg taggatcagc gactggggtg aagtcgtaac aaggtagccg taggggaacc  1440
```
E--> 106 tacggctgga tcac   → 1454
    107 (1454)

```
325 <210> SEQ ID NO: 5
326 <211> LENGTH: 729
327 <212> TYPE: DNA
328 <213> ORGANISM: crtW gene
330 <400> SEQUENCE: 5
331 atgagcgcac atgccctgcc caaggcagat ctgaccgcca ccagcctgat cgtctcgggc   60
333 ggcatcatcg ccgcgtggct ggccctgcat gtgcatgcgc tgtggtttct ggacgcggcg  120
335 gcgcatccca tcctggcgat cgcgaatttc ctggggctga cctggctgtc ggtcggtctg  180
337 ttcttcatcg cgcatgacgc gatgcacggg tcggtcgtgc cggggcgtcc gcgcggcaat  240
339 gcggcgatgg gccagctggt cctgtggctg tatgccggat tttcgtggcg caagatgatc  300
341 gtcaagcaca tggcccatca ccgccatacc ggaaccgacg acgaccccga tttcgaccat  360
343 ggcggcccgg tccgctggta cgcgcgcttc atcggcacct atttcggctg gcgcgagggg  420
345 ctgctgctgc ccgtcatcgt gacggtctat gcgctgatcc tggggatcg ctggatgtac  480
347 gtggtcttct ggccgctgcc gtcgatcctg gcgtcgatcc agctgttcgt gttcggcacc  540
349 tggctgccgc accgccccgg ccacgacgcg ttcccggacc gccataatgc gcggtcgtcg  600
351 cggatcagcg acccccgtgtc gctgctgacc tgctttcact ttggtggtta tcatcacgaa  660
353 caccacctgc acccgacggt gccttggtgg cgcctgccca gcacccgcac caagggggac  720
```
E--> 355 accgcatga   → 729
    356 (729)

```
359 <210> SEQ ID NO: 6                    ✓ OK
360 <211> LENGTH: (242)
361 <212> TYPE: PRT
362 <213> ORGANISM: crtW amino acid
364 <400> SEQUENCE: 6
365 Met Ser Ala His Ala Leu Pro Lys Ala Asp Leu Thr Ala Thr Ser Leu
```
E--> 366 1   (5)   10 (10)   15 (15)
```
368 Ile Val Ser Gly Gly Ile Ile Ala Ala Trp Leu Ala Leu His Val His
```
E--> 369   20 (20)   25 (25)   30 (30)
```
371 Ala Leu Trp Phe Leu Asp Ala Ala Ala His Pro Ile Leu Ala Ile Ala
```
E--> 372   35 (35)   40 (40)   45 L (45)
```
374 Asn Phe Leu Gly Leu Thr Trp Leu Ser Val Gly Leu Phe Phe Ile Ala
```
E--> 375   (50)   55 (55)   60 (60)
```
377 His Asp Ala Met His Gly Ser Val Val Pro Gly Arg Pro Arg Gly Asn
```
E--> 378 (65)   70 (70)   75 (75)   80
```
380 Ala Ala Met Gly Gln Leu Val Leu Trp Leu Tyr Ala Gly Phe Ser Trp
```
E--> 381   85 (85)   90 (90)   95 (95)
```
383 Arg Lys Met Ile Val Lys His Met Ala His His Arg His Thr Gly Thr
```
E--> 384 100 (100)   105 // 105   110 (110)
```
386 Asp Asp Asp Pro Asp Phe Asp His Gly Gly Pro Val Arg Trp Tyr Ala
```
E--> 387 (115)   120 (120)   125 (125)
```
389 Arg Phe Ile Gly Thr Tyr Phe Gly Trp Arg Glu Gly Leu Leu Leu Pro
```
       130   135   140

Handwritten annotation (right margin): Mis-aligned Amino Acid No. Pls See Item # 3 On error Summary Sheet.

RAW SEQUENCE LISTING   DATE: 10/14/2005
PATENT APPLICATION: US/10/551,508   TIME: 11:04:23

Input Set : N:\PANTU\10551508.raw.txt
Output Set: N:\CRF4\10142005\J551508.raw

```
E--> 390        130                     135                     140
     392 Val Ile Val Thr Val Tyr Ala Leu Ile Leu Gly Asp Arg Trp Met Tyr
E--> 393 145                  150 150               155  155            160      160
     395 Val Val Phe Trp Pro Leu Pro Ser Ile Leu Ala Ser Ile Gln Leu Phe
E--> 396                  165 165              170  170            175       175
     398 Val Phe Gly Thr Trp Leu Pro His Arg Pro Gly His Asp Ala Phe Pro
E--> 399              180 180             185 185            190      190
     401 Asp Arg His Asn Ala Arg Ser Ser Arg Ile Ser Asp Pro Val Ser Leu
E--> 402              195           200 200            205      205
     404 Leu Thr Cys Phe His Phe Gly Gly Tyr His His Glu His His Leu His
E--> 405          210              215 215            220  220
     407 Pro Thr Val Pro Trp Trp Arg Leu Pro Ser Thr Arg Thr Lys Gly Asp
E--> 408 225                  230 230            235  233            240       240
     410 Thr Ala
     414 <210> SEQ ID NO: 7
     415 <211> LENGTH: 489
     416 <212> TYPE: DNA
     417 <213> ORGANISM: crtZ gene
     419 <400> SEQUENCE: 7
     420 atgaccaatt tcctgatcgt cgtcgccacc gtgctggtga tggagttgac ggcctattcc    60
     422 gtccaccgtt ggatcatgca cggcccctg ggctggggct ggcacaagtc ccaccacgag   120
     424 gaacacgacc acgcgctgga aaagaacgac ctgtacggcc tggtctttgc ggtgatcgcc   180
     426 acggtgctgt tcacggtggg ctggatctgg gcgccggtcc tgtggtggat cgctttgggc   240
     428 atgaccgtct atgggctgat ctatttcgtc ctgcatgacg ggctggttca tcagcgctgg   300
     430 ccgttccgct atatcccgcg caagggctat gcccgccgcc tgtatcaggc ccaccgcctg   360
     432 caccacgcgg tcgagggacg cgaccattgc gtcagcttcg gcttcatcta tgcgccgccg   420
     434 gtcgacaagc tgaagcagga cctgaagacg tcgggcgtgc tgcgggccga ggcgcaggag   480
E--> 436 cgcacgtga                                                            489
     437 489
     440 <210> SEQ ID NO: 8
     441 <211> LENGTH: 162
     442 <212> TYPE: PRT
     443 <213> ORGANISM: crtZ amino acid
     445 <400> SEQUENCE: 8
     446 Met Thr Asn Phe Leu Ile Val Val Ala Thr Val Leu Val Met Glu Leu
E--> 447 1                   5                   10                  15
     449 Thr Ala Tyr Ser Val His Arg Trp Ile Met His Gly Pro Leu Gly Trp
E--> 450                     20                  25                  30
     452 Gly Trp His Lys Ser His His Glu Glu His Asp His Ala Leu Glu Lys
E--> 453                     35                  40                  45
     455 Asn Asp Leu Tyr Gly Leu Val Phe Ala Val Ile Ala Thr Val Leu Phe
E--> 456 50                  55                  60
     458 Thr Val Gly Trp Ile Trp Ala Pro Val Leu Trp Trp Ile Ala Leu Gly
E--> 459 65                  70                  75                  80
     461 Met Thr Val Tyr Gly Leu Ile Tyr Phe Val Leu His Asp Gly Leu Val
E--> 462                     85                  90                  95
     464 His Gln Arg Trp Pro Phe Arg Tyr Ile Pro Arg Lys Gly Tyr Ala Arg
E--> 465                     100                 105                 110
     467 Arg Leu Tyr Gln Ala His Arg Leu His His Ala Val Glu Gly Arg Asp
```

Same Error (handwritten annotation)

```
E--> 468          115          .              120                      125   .
     470 His Cys Val Ser Phe Gly Phe Ile Tyr Ala Pro Pro Val Asp Lys Leu
E--> 471          130                     135                  140
     473 Lys Gln Asp Leu Lys Thr Ser Gly Val Leu Arg Ala Glu Ala Gln Glu
E--> 474 145                   150                     155                        160
     476 Arg Thr
     527 <210> SEQ ID NO: 10
     528 <211> LENGTH: 386
     529 <212> TYPE: PRT
     530 <213> ORGANISM: crtY amino acid
     532 <400> SEQUENCE: 10
     533 Val Thr His Asp Val Leu Leu Ala Gly Ala Gly Leu Ala Asn Gly Leu
E--> 534   1                   5                      10                   15
     536 Ile Ala Leu Ala Leu Arg Ala Ala Arg Pro Asp Leu Arg Val Leu Leu
E--> 537          20                     25                      30
     539 Leu Asp His Ala Ala Gly Pro Ser Asp Gly His Thr Trp Ser Cys His
E--> 540          35                     40                      45
     542 Asp Pro Asp Leu Ser Pro His Trp Leu Ala Arg Leu Lys Pro Leu Arg
E--> 543          50                     55                      60
     545 Arg Ala Asn Trp Pro Asp Gln Glu Val Arg Phe Pro Arg His Ala Arg
E--> 546   65                   70                      75                   80
     548 Arg Leu Ala Thr Gly Tyr Gly Ser Leu Asp Gly Ala Ala Leu Ala Asp
E--> 549               85                     90                   95
     551 Ala Val Ala Arg Ser Gly Ala Glu Ile Arg Trp Asn Ser Asp Ile Ala
E--> 552          100                    105                    110
     554 Leu Leu Asp Glu Gln Gly Ala Thr Leu Ser Cys Gly Thr Arg Ile Glu
E--> 555          115         .          120                    125
     557 Ala Gly Ala Val Leu Asp Gly Arg Gly Ala Gln Pro Ser Arg His Leu
E--> 558   .      130                    135                  140
     560 Thr Val Gly Phe Gln Lys Phe Val Gly Val Glu Ile Glu Thr Asp Cys
E--> 561 145                   150                     155                   160
     563 Pro His Gly Val Pro Arg Pro Met Ile Met Asp Ala Thr Val Thr Gln
E--> 564               165   .             170                  175
     566 Gln Asp Gly Tyr Arg Phe Ile Tyr Leu Leu Pro Phe Ser Pro Thr Arg
E--> 567          180                    185                    190
     569 Ile Leu Ile Glu Asp Thr Arg Tyr Ser Asp Gly Gly Asn Leu Asp Asp
E--> 570          195                    200                    205
     572 Asp Ala Leu Ala Ala Ala Ser His Asp Tyr Ala Arg Gln Gln Gly Trp
E--> 573          210                    215                    220
     575 Thr Gly Ala Glu Val Arg Arg Glu Arg Gly Ile Leu Pro Ile Ala Leu
E--> 576 225                   230                     235                  240
     578 Ala His Asp Ala Ala Gly Phe Trp Ala Asp His Ala Glu Gly Pro Val
E--> 579               245                    250                   255
     581 Pro Val Gly Leu Arg Ala Gly Phe Phe His Pro Val Thr Gly Tyr Ser
E--> 582   .           260                    265                    270
     584 Leu Pro Tyr Ala Ala Gln Val Ala Asp Val Val Ala Gly Leu Ser Gly
E--> 585          275                    280                    285
     587 Pro Pro Gly Thr Asp Ala Leu Arg Gly Ala Ile Arg Asp Tyr Ala Ile
E--> 588          290                    295                    300
```

Same Error Mis-aligned

RAW SEQUENCE LISTING
PATENT APPLICATION: US/10/551,508

DATE: 10/14/2005
TIME: 11:04:23

Input Set : N:\FANTU\10551508.raw.txt
Output Set: N:\CRF4\10142005\J551508.raw

```
      590 Asp Arg Ala Arg Arg Asp Arg Phe Leu Arg Leu Leu Asn Arg Met Leu
E--> 591 305                     310                     315                          320
      593 Phe Arg Gly Cys Ala Pro Asp Arg Arg Tyr Thr Leu Leu Gln Arg Phe
E--> 594                         325·                    330                     335
      596 Tyr Arg Met Pro His Gly Leu Ile Glu Arg Phe Tyr Ala Gly Arg Leu
E--> 597 ·             340                     345                     350
      599 Ser Val Ala Asp Gln Leu Arg Ile Val Thr Gly Lys Pro Pro Ile Pro
E--> 600           355                     360                     365
      602 Leu Gly Thr Ala Ile Arg Cys Leu Pro Glu Arg Pro Leu Leu Lys Glu
E--> 603       370                     375                     380
      605 Asn Ala
      606 385
      609 <210> SEQ ID NO: 11
      610 <211> LENGTH: 1506
      611 <212> TYPE: DNA
      612 <213> ORGANISM: crtI gene
      614 <400> SEQUENCE: 11
      615 atgaacgccc attcgcccgc ggccaagacc gccatcgtga tcggcgcagg ctttggcggg    60
     ·617 ctggccctgg ccatccgcct gcagtccgcg ggcatcgcca ccaccctggt cgaggcccgg   120
      619 gacaagcccg gcgggcgcgc ctatgtctgg cacgatcagg gccatgtctt cgacgcgggc   180
     ·621 ccgaccgtca tcaccgaccc cgatgcgctc aaggagctgt gggcgctgac cgggcaggac   240
      623 atggcgcgcg acgtgacgct gatgccggtg tcgcccttct atcgactgat gtggccgggc   300
      625 gggaaggtct tcgattacgt gaacgaggcc gatcagctgg agcgccagat cgcccagttc   360
      627 aacccggacg acctggaagg ataccgccgc ttccgtgatt acgcggagga ggtgtatcag·   420
      629 gagggctacg tcaagctggg caccgtgccc ttcctcaagc tgggccagat gctcaaggcc   480
      631 gcgcccgcgc tgatgaagct ggaggcctat aagtccgtcc atgccaaggt cgcgaccttc   540
      633 atcaaggacc cctatctgcg gcaggcgttt tcgtatcaca cgctgctggt gggcgggaat   600
      635 cccttctcga ccagctcgat ctatgcgctg atccacgcgc tggagcggcg cggcggggtc   660
      637 tggttcgcca agggcggcac caaccagctg gtcgcgggca tggtcgcgct gttcgaacgg   720
     .639 cttggcggcc ·agatgatgct gaacgccaag gtcgcccgga tcgagaccga gggcgcgcgg   780
      641 accacgggcg tcaccctggc ggacgggcgg tctttaaggg ccgacatggt .cgccagcaac   840
      643 ggcgacgtca tgcacaacta tcgcgacctg ctgggccaca cggcccgcgg gcagagccgc   900
      645 gcgaaatcgc tggaccgcaa gcgctggtcc atgtcgttgt cgtgctgca tttcggtctg   960
      647 cgcgaggcgc ccaaggacat cgcgcatcac accatcctgt tcggcccccg ctacagggag  1020
      649 ctggtcaacg agatcttcaa gggccccgaag ctggccgagg atttctcgct gtacctgcat  1080
      651 tcgccctgca cgaccgatcc ggacatggcg cctccgggca tgtccacgca ttacgtgctg  1140
      653 gcccccgtgc cgcatctggg ccgcgccgag atcgattggg cggtcgaggg gccgcgctat ·1200
      655 gccgaccgca tcctggcgtc cctggaggag cggctgatcc cgaacctgcg cgccaacctg  1260
      657 accacgacgc gcatcttcac gcccgccgat ttcgccagcg aactgaacgc catcacggc  1320
      659 agcgccttct cggtcgagcc gatcctgacg caatccgcgt ggttccggcc gcacaaccgc  1380
      661 gacaagacga tccgcaactt ctatctggtc ggcgcgggca cccatccggg cgcgggcatt  1440
      663 ccgggcgtcg tgggctcggc caaggccacg gcccaggtga tgctgtccga cctggcgggc  1500
E--> 665 gcatga                                                             ⌐1506
      666 1506⌐
      669 <210> SEQ ID NO: 12
      670 <211> LENGTH: 501
      671 <212> TYPE: PRT
      672 <213> ORGANISM: crtI amino acid
      674 <400> SEQUENCE: 12
```

RAW SEQUENCE LISTING   DATE: 10/14/2005
PATENT APPLICATION: US/10/551,508   TIME: 11:04:23

Input Set : N:\FANTU\10551508.raw.txt
Output Set: N:\CRF4\10142005\J551508.raw

```
      675 Met Asn Ala His Ser Pro Ala Ala Lys Thr Ala Ile Val Ile Gly Ala
B--> 676  1               5                   10                      15
      678 Gly Phe Gly Gly Leu Ala Leu Ala Ile Arg Leu Gln Ser Ala Gly Ile
B--> 679          20                  25                      30
      681 Ala Thr Thr Leu Val Glu Ala Arg Asp Lys Pro Gly Gly Arg Ala Tyr
B--> 682          35                  40                      45
      684 Val Trp His Asp Gln Gly His Val Phe Asp Ala Gly Pro Thr Val Ile
B--> 685          50                  55                      60
      687 Thr Asp Pro Asp Ala Leu Lys Glu Leu Trp Ala Leu Thr Gly Gln Asp
B--> 688 65                  70                      75                  80
      690 Met Ala Arg Asp Val Thr Leu Met Pro Val Ser Pro Phe Tyr Arg Leu
B--> 691          85                  90                      95
      693 Met Trp Pro Gly Gly Lys Val Phe Asp Tyr Val Asn Glu Ala Asp Gln
B--> 694          100                 105                     110
      696 Leu Glu Arg Gln Ile Ala Gln Phe Asn Pro Asp Asp Leu Glu Gly Tyr
B--> 697          115                 120                     125
      699 Arg Arg Phe Arg Asp Tyr Ala Glu Glu Val Tyr Gln Glu Gly Tyr Val
B--> 700          130                 135                     140
      702 Lys Leu Gly Thr Val Pro Phe Leu Lys Leu Gly Gln Met Leu Lys Ala
B--> 703 145                 150                     155                 160
      705 Ala Pro Ala Leu Met Lys Leu Glu Ala Tyr Lys Ser Val His Ala Lys
B--> 706          165                 170                     175
      708 Val Ala Thr Phe Ile Lys Asp Pro Tyr Leu Arg Gln Ala Phe Ser Tyr
B--> 709          180                 185                     190
      711 His Thr Leu Leu Val Gly Gly Asn Pro Phe Ser Thr Ser Ser Ile Tyr
B--> 712          195                 200                     205
      714 Ala Leu Ile His Ala Leu Glu Arg Arg Gly Gly Val Trp Phe Ala Lys
B--> 715          210                 215                     220
      717 Gly Gly Thr Asn Gln Leu Val Ala Gly Met Val Ala Leu Phe Glu Arg
B--> 718 225                 230                     235                 240
      720 Leu Gly Gly Gln Met Met Leu Asn Ala Lys Val Ala Arg Ile Glu Thr
B--> 721          245                 250                     255
      723 Glu Gly Ala Arg Thr Thr Gly Val Thr Leu Ala Asp Gly Arg Ser Leu
B--> 724          260                 265                     270
      726 Arg Ala Asp Met Val Ala Ser Asn Gly Asp Val Met His Asn Tyr Arg
B--> 727          275                 280                     285
      729 Asp Leu Leu Gly His Thr Ala Arg Gly Gln Ser Arg Ala Lys Ser Leu
B--> 730          290                 295                     300
      732 Asp Arg Lys Arg Trp Ser Met Ser Leu Phe Val Leu His Phe Gly Leu
B--> 733 305                 310                     315                 320
      735 Arg Glu Ala Pro Lys Asp Ile Ala His His Thr Ile Leu Phe Gly Pro
B--> 736          325                 330                     335
      738 Arg Tyr Arg Glu Leu Val Asn Glu Ile Phe Lys Gly Pro Lys Leu Ala
B--> 739          340                 345                     350
      741 Glu Asp Phe Ser Leu Tyr Leu His Ser Pro Cys Thr Thr Asp Pro Asp
B--> 742          355                 360                     365
      744 Met Ala Pro Pro Gly Met Ser Thr His Tyr Val Leu Ala Pro Val Pro
B--> 745          370                 375                     380
      747 His Leu Gly Arg Ala Glu Ile Asp Trp Ala Val Glu Gly Pro Arg Tyr
```

RAW SEQUENCE LISTING   DATE: 10/14/2005
PATENT APPLICATION: US/10/551,508   TIME: 11:04:23

Input Set : N:\FANTU\10551508.raw.txt
Output Set: N:\CRF4\10142005\J551508.raw

```
E--> 748 385                    390                    395                    400
     750 Ala Asp Arg Ile Leu Ala Ser Leu Glu Glu Arg Leu Ile Pro Asn Leu
E--> 751              405                    410                    415
     753 Arg Ala Asn Leu Thr Thr Thr Arg Ile Phe Thr Pro Ala Asp Phe Ala
E--> 754              420                    425                    430
     756 Ser Glu Leu Asn Ala His His Gly Ser Ala Phe Ser Val Glu Pro Ile
E--> 757              435                    440                    445
     759 Leu Thr Gln Ser Ala Trp Phe Arg Pro His Asn Arg Asp Lys Thr Ile
E--> 760       450                    455                    460
     762 Arg Asn Phe Tyr Leu Val Gly Ala Gly Thr His Pro Gly Ala Gly Ile
E--> 763 465                    470                    475                    480
     765 Pro Gly Val Val Gly Ser Ala Lys Ala Thr Ala Gln Val Met Leu Ser
E--> 766              485                    490                    495
     768 Asp Leu Ala Gly Ala
E--> 769              500
     772 <210> SEQ ID NO: 13
     773 <211> LENGTH: 915
     774 <212> TYPE: DNA
     775 <213> ORGANISM: crtB gene
     777 <400> SEQUENCE: 13
     778 atgagcgatc tggtcctgac ctcgaccgag gcgatcaccc aagggtcgca aagctttgcc     60
     780 acggcggcca agctgatgcc gccgggcatc cgcgacgaca cggtgatgct ctatgcctgg    120
     782 tgccgccacg cggatgacgt gatcgacggt caggccctgg gcagccgccc cgaggcggtg    180
     784 aacgacccgc aggcgcggct ggacggcctg cgcgtcgaca cgctggcggc cctgcagggc    240
     786 gacggtccgg tgaccccgcc ctttgccgcg ctgcgcgcgg tggcgcggcg gcatgatttc    300
     788 ccgcaggcct ggcccatgga cctgatcgaa ggcttcgcga tggatgtcga ggcgcgcgac    360
     790 tatcgcacgc tggatgacgt gctggaatat tcctatcacg tcgcaggcat cgtcggcgtg    420
     792 atgatggccc gcgtgatggg cgtgcgcgac gatcctgtcc tggaccgcgc ctgcgacctg    480
     794 gggctggcgt tccagctgac caacatcgcg cgcgacgtga tcgacgatgc gcgcatcggg    540
     796 cggtgctatc tgccggggga ctggctggac caggcgggcg cgcggatcga cgggccggtg    600
     798 ccgtcgccgg agctgtacac agtgatcctc cggctgttgg atgaggcgga accctattac    660
     800 gcgtcggcgc gggtggggtct ggcggatctg ccaccgcgct cgcgcctggtc catcgccgcc    720
     802 gcgctacgga tctatcgcgc catcgggctg cgcatccgca agagcgggcc gcaggcctat    780
     804 cgccagcgga tcagcacgtc caaggctgcc aagatcggcc tgctgggcgt cggggggctgg    840
     806 gatgtcgcgc gatcacgcct gccggggggcg ggcgtgtcgc ggcagggcct ctggacccgg    900
E--> 808 ccgcatcacg tctag    915
     809 915
     812 <210> SEQ ID NO: 14
     813 <211> LENGTH: 304
     814 <212> TYPE: PRT
     815 <213> ORGANISM: crtB amino acid
     817 <400> SEQUENCE: 14
     818 Met Ser Asp Leu Val Leu Thr Ser Thr Glu Ala Ile Thr Gln Gly Ser
E--> 819  1                    5                    10                    15
     821 Gln Ser Phe Ala Thr Ala Ala Lys Leu Met Pro Pro Gly Ile Arg Asp
E--> 822              20                    25                    30
     824 Asp Thr Val Met Leu Tyr Ala Trp Cys Arg His Ala Asp Asp Val Ile
E--> 825       35                    40                    45
     827 Asp Gly Gln Ala Leu Gly Ser Arg Pro Glu Ala Val Asn Asp Pro Gln
```

```
E--> 828        50                    55                    60
     830 Ala Arg Leu Asp Gly Leu Arg Val Asp Thr Leu Ala Ala Leu Gln Gly
E--> 831 65                    70                    75                    80
     833 Asp Gly Pro Val Thr Pro Pro Phe Ala Ala Leu Arg Ala Val Ala Arg
E--> 834                   85                    90                .      95
     836 Arg His Asp Phe Pro Gln Ala Trp Pro Met Asp Leu Ile Glu Gly Phe
E--> 837               100                   105                   110
     839 Ala Met Asp Val Glu Ala Arg Asp Tyr Arg Thr Leu Asp Asp Val Leu
E--> 840           115                   120                   125
     842 Glu Tyr Ser Tyr His Val Ala Gly Ile Val Gly Val Met Met Ala Arg
E--> 843       130                   135                   140
     845 Val Met Gly Val Arg Asp Asp Pro Val Leu Asp Arg Ala Cys Asp Leu
E--> 846 145                   150                   155                   160
     848 Gly Leu Ala Phe Gln Leu Thr Asn Ile Ala Arg Asp Val Ile Asp Asp
E--> 849                   165                   170             175
     851 Ala Arg Ile Gly Arg Cys Tyr Leu Pro Gly Asp Trp Leu Asp Gln Ala
E--> 852               180                   185                   190
     854 Gly Ala Arg Ile Asp Gly Pro Val Pro Ser Pro Glu Leu Tyr Thr Val
E--> 855           195                   200                   205
     857 Ile Leu Arg Leu Leu Asp Glu Ala Glu Pro Tyr Tyr Ala Ser Ala Arg
E--> 858       210                   215                   220
     860 Val Gly Leu Ala Asp Leu Pro Pro Arg Cys Ala Trp Ser Ile Ala Ala
E--> 861 225                   230                   235                   240
     863 Ala Leu Arg Ile Tyr Arg Ala Ile Gly Leu Arg Ile Arg Lys Ser Gly
E--> 864                   245                   250             255
     866 Pro Gln Ala Tyr Arg Gln Arg Ile Ser Thr Ser Lys Ala Ala Lys Ile
E--> 867               260                   265                   270
     869 Gly Leu Leu Gly Val Gly Gly Trp Asp Val Ala Arg Ser Arg Leu Pro
E--> 870           275                   280                   285
     872 Gly Ala Gly Val Ser Arg Gln Gly Leu Trp Thr Arg Pro His His Val
E--> 873       290                   295                   300
     915 <210> SEQ ID NO: 16
     916 <211> LENGTH: 293
     917 <212> TYPE: PRT
     918 <213> ORGANISM: crtE amino acid
     920 <400> SEQUENCE: 16
     921 Met Arg Arg Asp Val Asn Pro Ile His Ala Thr Leu Leu Gln Thr Arg
E--> 922 1                     5                     10                    15
     924 Leu Glu Glu Ile Ala Gln Gly Phe Gly Ala Val Ser Gln Pro Leu Gly
E--> 925               20                    25                    30
     927 Ala Ala Met Ser His Gly Ala Leu Ser Ser Gly Arg Arg Phe Arg Gly
E--> 928           35                    40                    45
     930 Met Leu Met Leu Leu Ala Ala Glu Ala Ser Gly Gly Val Cys Asp Thr
E--> 931 50                    55                    60
     933 Ile Val Asp Ala Ala Cys Ala Val Glu Met Val His Ala Ala Ser Leu
E--> 934 65                    70                    75                    80
     936 Ile Phe Asp Asp Leu Pro Cys Met Asp Asp Ala Gly Leu Arg Arg Gly
E--> 937               85                    90                    95
     939 Arg Pro Ala Thr His Val Ala His Gly Glu Ser Arg Ala Val Leu Gly
```

```
E--> 940            100                   105                  110
     942 Gly Ile Ala Leu Ile Thr Glu Ala Met Ala Leu Leu Ala Gly Ala Arg
E--> 943            115                   120                  125
     945 Gly Ala Ser Gly Thr Val Arg Ala Gln Leu Val Arg Ile Leu Ser Arg
E--> 946        130               135                 140
     948 Ser Leu Gly Pro Gln Gly Leu Cys Ala Gly Gln Asp Leu Asp Leu His
E--> 949 145                150                   155                160
     951 Ala Ala Lys Asn Gly Ala Gly Val Glu Gln Glu Gln Asp Leu Lys Thr
E--> 952            165                   170                 175
     954 Gly Val Leu Phe Ile Ala Gly Leu Glu Met Leu Ala Val Ile Lys Glu
E--> 955            180                   185                  190
     957 Phe Asp Ala Glu Glu Gln Thr Gln Met Ile Asp Phe Gly Arg Gln Leu
E--> 958            195                   200                  205
     960 Gly Arg Val Phe Gln Ser Tyr Asp Asp Leu Leu Asp Val Val Gly Asp
E--> 961        210               215                   220
     963 Gln Ala Ala Leu Gly Lys Asp Thr Gly Arg Asp Ala Ala Ala Pro Gly
E--> 964 225                230                   235                240
     966 Pro Arg Arg Gly Leu Leu Ala Val Ser Asp Leu Gln Asn Val Ser Arg
E--> 967            245                   250                 255
     969 His Tyr Glu Ala Ser Arg Ala Gln Leu Asp Ala Met Leu Arg Ser Lys
E--> 970            260                   265                  270
     972 Arg Leu Gln Ala Pro Glu Ile Ala Ala Leu Leu Glu Arg Val Leu Pro
E--> 973            275                   280                  285
     975 Tyr Ala Ala Arg Ala
     976     290
```

same error

RAW SEQUENCE LISTING ERROR SUMMARY   DATE: 10/14/2005
PATENT APPLICATION: US/10/551,508   TIME: 11:04:24

Input Set : N:\PANTU\10551508.raw.txt
Output Set: N:\CRF4\10142005\J551508.raw

Invalid Line Length:
The rules require that a line not exceed 72 characters in length. This includes spaces.

Seq#:1; Line(s) 7,8

VERIFICATION SUMMARY   DATE: 10/14/2005
PATENT APPLICATION: US/10/551,508   TIME: 11:04:24

Input Set : N:\PANTU\10551508.raw.txt
Output Set: N:\CRF4\10142005\J551508.raw

L:13 M:270 C: Current Application Number differs, Replaced Current Application No
L:13 M:271 C: Current Filing Date differs, Replaced Current Filing Date
L:106 M:254 E: No. of Bases conflict, LENGTH:Input:0 Counted:1454 SEQ:3
L:355 M:254 E: No. of Bases conflict, LENGTH:Input:0 Counted:729 SEQ:5
L:366 M:332 E: (32) Invalid/Missing Amino Acid Numbering, SEQ ID:6
M:332 Repeated in SeqNo=6
L:436 M:254 E: No. of Bases conflict, LENGTH:Input:0 Counted:489 SEQ:7
L:447 M:332 E: (32) Invalid/Missing Amino Acid Numbering, SEQ ID:8 -
M:332 Repeated in SeqNo=8
L:534 M:332 E: (32) Invalid/Missing Amino Acid Numbering, SEQ ID:10
M:332 Repeated in SeqNo=10
L:665 M:254 E: No. of Bases conflict, LENGTH:Input:0 Counted:1506 SEQ:11
L:676 M:332 E: (32) Invalid/Missing Amino Acid Numbering, SEQ ID:12
M:332 Repeated in SeqNo=12
L:808 M:254 E: No. of Bases conflict, LENGTH:Input:0 Counted:915 SEQ:13
L:819 M:332 E: (32) Invalid/Missing Amino Acid Numbering, SEQ ID:14
M:332 Repeated in SeqNo=14
L:922 M:332 E: (32) Invalid/Missing Amino Acid Numbering, SEQ ID:16
M:332 Repeated in SeqNo=16